National Defence / Défense nationale

UNCLASSIFIED

UNLIMITED DISTRIBUTION

DRES

SUFFIELD MEMORANDUM

NO. 1239

INFERENCE OF PARTICLE SIZE DISTRIBUTION FUNCTION OF AEROSOL CLOUDS FROM LIGHT SCATTERING MEASUREMENTS (U)

by

E. Yee

PCN 351SD

August 1988

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD, RALSTON, ALBERTA

DEFENSO PER SCIENTIAM

Canada

UNCLASSIFIED

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD

RALSTON ALBERTA

SUFFIELD MEMORANDUM NO. 1239

INFERENCE OF PARTICLE SIZE DISTRIBUTION FUNCTION OF AEROSOL CLOUDS FROM LIGHT SCATTERING MEASUREMENTS (U)

by

E. Yee

PCN 351SD

UNCLASSIFIED

UNCLASSIFIED

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD

RALSTON ALBERTA

SUFFIELD MEMORANDUM NO. 1239

INFERENCE OF PARTICLE SIZE DISTRIBUTION FUNCTION OF AEROSOL CLOUDS FROM LIGHT SCATTERING MEASUREMENTS (U)

by

E. Yee

ABSTRACT

Perhaps the most versatile and efficient method for inferring the particle size distribution function (PSDF) of aerosol clouds from remote light scattering measurements is the constrained linear inversion procedure due to Philips and Twomey. However, conventional numerical implementations of this procedure are subject to the following two problems: (1) an appropriate discrete approximation must be chosen for the PSDF which adequately achieves the correct balance between the conflicting requirements of resolution of detail in the PSDF and of efficiency in the computation of the solution and (2) a proper value must be selected for the Lagrange multiplier (smoothing parameter) that adequately reflects the tradeoff between the fidelity to the observed optical data and the smoothness of the solution. Consequently, an adequate recovery of the PSDF, based on the constrained linear inversion procedure, is usually achieved only after a certain amount of tedious preliminary exploratory analysis.

An alternative implementation of the constrained linear inversion procedure is presented which overcomes the problems associated with the conventional implementation. Firstly, an explicit analytical (continuous) representation for the solution of the constrained linear inversion procedure is developed which obviates the need to obtain a discrete approximation for the PSDF. Secondly, an objective procedure, based on the principle of generalized cross-validation, is utilized for the selection of the proper value for the Lagrange multiplier. Taken together, these two developments provide the basis for an objective, fully automated implementation of the constrained linear inversion technique. Numerical examples of PSDF inversions, obtained using the proposed automatic retrieval algorithm, are presented for synthetic aerosol optical extinction and scattering data.

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD

RALSTON ALBERTA

SUFFIELD MEMORANDUM NO. 1239

INFERENCE OF PARTICLE SIZE DISTRIBUTION FUNCTION OF AEROSOL CLOUDS FROM LIGHT SCATTERING MEASUREMENTS (U)

by

E. Yee

INTRODUCTION

1. The determination of the distribution of particle sizes in a polydispersed aerosol cloud, based on a remote sensing methodology, does not permit a direct measurement of the size spectra. Rather, the information concerning the particle size distribution function (PSDF) must be inferred from the light scattering data (viz., measurements of the way the aerosol cloud scatters electromagnetic radiation) which provide either the spectral extinction as a function of the incident wavelength and/or the scattered light intensity as a function of the scattering angle. Given some combination of these measurements, the recovery of the PSDF of the aerosol cloud is, in essence, an inverse or indirect problem.

2. The retrieval of the aerosol particle size spectra from a finite set of imprecise optical scattering data is an inherently ill-posed problem in the sense that the solution is both nonunique and unstable (viz., the solution does not depend continuously on the data). In view of the fact that a knowledge of the PSDF would aid in the understanding of the various physical and chemical mechanisms and processes which are responsible for the aerosol microstructure and would permit the development of more efficient and reliable particle diagnostic methods for the evaluation and assessment of airborne particulate hazards, a considerable research effort has been directed to the development of algorithms for the reconstruction of the aerosol size distribution from optical scattering data.

3. The earliest attempt at the specification of a PSDF retrieval algorithm can be attributed to Yamamoto and Tanaka [1], who formulated a numerical inversion algorithm based on the constrained linear inversion procedure originally developed by Philips [2] and Twomey [3] in the United States and, independently, by Tikhonov [4] in the Soviet Union. The method of constrained linear inversion has subsequently been applied to the inference of particle size distributions utilizing a myriad of light scattering measurements, including

optical depth data obtained as a function of wavelength [5], backscattered radiance measured as a function of wavelength [6], a combination of optical extinction determined as a function of wavelength and scattered light intensity determined as a function of scattering angle [7], and spectral light distribution and polarization distribution obtained as a function of wavelength and scattering angle [8].

4. Without a doubt, the constrained linear inversion procedure, which utilizes the Philips and Twomey second derivative smoothing constraint, is the most popular procedure for the inference of aerosol size distributions. However, in the implementation of this procedure, it is necessary to adopt a discrete representation for the PSDF $f(r)$ as well as a difference approximation for the second derivative smoothing functional. Almost invariably, $f(r)$ is approximated by a piece-wise constant representation whereby the size domain (range in radii r of the particles) is partitioned into a finite number of intervals over which $f(r)$ is assumed to be constant. This discretization leads to the following problem: there is no objective procedure for selecting the correct number of discrete segments to employ in the approximation of $f(r)$. Ideally, the number of segments must be large enough so that the resulting binning does not lead to a loss of detail in the PSDF and at the same time not so large that the resulting discretized problem cannot be solved economically on a computer. Moreover, the constrained linear inversion solution depends critically on the selection of a proper value for the Lagrange multiplier (smoothing parameter) γ that determines the degree of smoothing to impose on the solution. Again, there is no objective method for the selection of an appropriate value for γ. For the most part, attempts at a resolution to this problem have been based on trial-and-error procedures and/or on empirical rules derived either from computational experiments or from a sensitivity analysis [7,9].

5. In view of the problems associated with the application of the constrained linear inversion procedure, Curry and Kiech [10,11] have proposed two alternative procedures for the retrieval of the PSDF from light scattering data. The first procedure, referred to as the constrained eigenfunction expansion, expresses the PSDF in terms of the eigenfunctions of a covariance operator constructed from the Mie scattering kernels. The coefficients in this expansion are determined so that the recovered PSDF is as close to some trial solution (function) as possible while maintaining the required consistency with the noisy data. The constrained eigenfunction expansion procedure is actually a generalization of an analytic inversion procedure developed by Capps, Henning and Hess [12]. It is important to note that although the constrained eigenfunction expansion method does not require a discrete approximation for the PSDF, it still nevertheless requires the proper selection of the value of the Lagrange multiplier γ. Curry and Kiech proposed that γ be found by application of the Residual Relative Variance method, a process which leads to a somewhat complex double iterative computational sequence for the solution. The second procedure, designated as the nonlinear regression method, is precisely as the name implies—the PSDF is postulated to possess a particular functional form (e.g., log-normal) and the set of parameters associated with this form is then estimated directly from the data based on the least-squares method. This procedure can yield very good results provided the proper parametric model is chosen for the PSDF and the correct number of parameters are utilized. However, it

should be emphasized that the nonlinear regression method can also lead to very poor results if the latter two criteria are not adequately satisfied. Indeed, in view of the fact that, in most cases, there is no a priori information on the correct parametric form to use for the PSDF, the nonlinear regression procedure cannot be recommended for general use (i.e., for "blind" inversion). Moreover, the nonlinear regression method is computationally intensive due to the nonlinear occurrence of the parameters in the assumed parametric (i.e., model) size distribution function.

6. Despite the introduction of the constrained eigenfunction expansion and the nonlinear regression procedures, the constrained linear inversion method still remains by far the most widely employed method for inferring particle size distributions from optical data. This is because the latter method only requires that the unknown solution satisfy a weak a prior constraint (i.e., a smoothness constraint) unlike the constrained eigenfunction expansion and nonlinear regression methods. Moreover, the constrained linear inversion method can lead to very good results provided a certain amount of careful preliminary exploratory analysis is undertaken to select the proper discrete representation for the PSDF and the proper value for the smoothing parameter γ. In view of this, it would be desirable to develop a fully automated implementation of the constrained linear inversion procedure which (1) does not require a discrete representation for the PSDF, thus obviating the need to select the proper number of discrete segments with which to divide the size domain, and (2) incorporates an objective method for the selection of the appropriate value for γ.

7. The major feature that distinguish this work from that of previous authors on the implementation of the constrained linear inversion procedure is the emphasis on the development of a completely automatic implementation of the procedure that requires little or no user intervention. Consequently, the purpose of this paper is to show how to achieve a completely automatic implementation of the constrained linear inversion procedure. To this end, it is first shown that, for a fixed value of the smoothing parameter γ, an analytical representation for the solution of the constrained linear inversion method can be obtained by utilizing standard techniques from the theory of abstract Hilbert spaces. This representation expresses the PSDF as a continuously differentiable function of the radius r and, as such, obviates the requirement for the selection of an appropriate discrete representation as is required in the usual application of the constrained linear inversion method. Consequently, the problems stemming from the proper discretization of the PSDF are no longer an issue. Next, it is shown how the principle of generalized cross-validation can be applied to provide an objective procedure for the selection of an appropriate value of the Lagrange multiplier (i.e., smoothing parameter). In this way, the second major problem associated with the application of the constrained linear inversion procedure can be resolved. Indeed, the coupling of the analytic representation of the PSDF with the cross-validated selection of the Lagrange multiplier, provides the basis for a fully automatic implementation of the constrained linear inversion method with no need for a preliminary (exploratory) analysis on the user's part.

MATHEMATICAL FORMULATION OF THE PROBLEM

8. Electromagnetic theory dictates that optical scattering data from a polydispersed aerosol cloud consisting of an ensemble of homogeneous spherical particles is related to the particle size distribution function through a Fredholm integral equation of the first kind [13,14]

$$s(x) = \int_{\Omega} K(x,r) f(r)\, dr, \tag{1a}$$

provided that the aerosol cloud is optically thin so that only single scattering occurs and that the particles comprising the cloud are randomly separated so that only incoherent scattering occurs (i.e., there is no systematic phase relation between the radiation scattered by the individual particles). In eq. (1a), $s(x)$ denotes the theoretical (model) optical data (e.g., optical extinction, spectral turbidity, scattered light intensity, etc.) determined as a function of the wavelength $x \equiv \lambda$ and/or the scattering angle $x \equiv \theta$. The kernel function $K(x,r)$, corresponding to the datum $s(x)$, represents the scattering, backscatter, or extinction cross sections for a spherical particle of radius r. This function, which can be evaluated using Mie theory, is also dependent on the complex index of refraction m of the particle. In the following treatment, it is assumed that m is known. Finally, $f(r)$ in eq. (1a) denotes the particle size distribution function and $\Omega \equiv [R_1, R_2]$ specifies the radii limits R_1 and R_2 (i.e., size domain) for $f(r)$.

9. A particular experiment will yield M numerical values of $s(x)$ (e.g., obtained for a finite and discrete set of measurement variables $\{x_i\}_{i=1}^{M}$) which are contaminated with noise. Consequently, the observational model relating the measured data D_i to the model data $s(x_i)$ is given by

$$D_i = s(x_i) + \epsilon_i, \qquad i = 1,2,\ldots,M, \tag{1b}$$

where ϵ_i is assumed to be a zero-mean measurement error with variance σ_i^2. The inverse problem associated with eqs. (1a) and (1b) can now be stated as follows: given a finite set of inaccurate optical scattering data $\{D_i\}_{i=1}^{M}$ and the errors $\{\sigma_i\}_{i=1}^{M}$ in their measurement, determine the PSDF $f(r)$.

10. The most important property to note with regard to this inverse problem is that it is inherently ill-posed. This translates into the observation that a small random perturbation on the optical data $s(x)$ will induce unacceptably large variations on the solution $f(r)$. Consequently, a practical solution to the problem can only be obtained if the ill-posedness is directly confronted and addressed. If this is not done, the computation of the solution from practical data (cf. eq. (1b)) will invariably lead to wild oscillations in the PSDF. These oscillations, which are merely artifacts in the solution that have no direct physical significance, are the manifestations of the ill-posed nature of the problem. To formulate a well-posed problem, Philips [2] and Twomey [3] have proposed that a regularized solution (or "quasi-solution") $f_\gamma(r)$ be obtained by minimizing a smoothing functional of the form

$$R_\gamma(f) = R_{\mathrm{LS}}(f) + \gamma R_{\mathrm{S}}(f), \tag{2a}$$

where

$$R_{\rm LS}(f) = \frac{1}{M}\sum_{i=1}^{M}\frac{\left(D_i - s(x_i)\right)^2}{\sigma_i^2} \tag{2b}$$

is the usual weighted least-squares discrepancy (misfit) function with the i-th weight chosen to be the inverse of the variance σ_i^2 associated with the i-th datum D_i and

$$R_{\rm S}(f) = \int_{\Omega}\left(\frac{d^2 f}{dr^2}\right)^2 dr \tag{2c}$$

is a stabilizing functional (penalty term) which measures the degree of smoothness (regularity) in the unknown PSDF. Note that $R_{\rm S}(f)$ discriminates against steep curvatures in $f(r)$ and, in this sense at least, the PSDF obtained by minimizing $R_{\rm S}(f)$ subject to the constraint that $f(r)$ reproduce the data to within some expected accuracy, must generate the "flattest" solution consistent with the available data.

11. In eq. (2a), γ is a non-negative regularization parameter (Lagrange multiplier). The smoothing functional is composed of the linear combination of the data misfit $R_{\rm LS}(f)$ and the roughness measure $R_{\rm S}(f)$, with the latter weighted by the regularization parameter γ. This parameter must be chosen to reflect the degree of smoothing to be imposed on the required solution and, indeed, as γ increases, the roughness term in eq. (2a) is given an increasing significance (i.e., weight) at the expense of the misfit term. Ideally, the regularization parameter must be chosen to reflect some desirable compromise between the fidelity with which the PSDF regenerates the data as measured by the misfit term and the smoothness desired in the PSDF as measured by the roughness term. It should perhaps be noted that the choice of γ is equivalent to the choice of the misfit tolerance T, which reflects the degree of consistency between the theory and observations in the form $R_{\rm LS}(f) \le T$. Of course, T (or, equivalently γ) must be chosen to adequately reflect the magnitude of the noise perturbing the data.

12. The problem posed in eq. (2a) is usually solved [2,3,15] by dividing the size domain $\Omega = [R_1, R_2]$ into N non-intersecting equal-sized intervals $\Delta r_j = r_j - r_{j-1}$ $(j = 1, 2, \ldots, N)$ with $f(r)$ assumed to be constant over each interval, viz.

$$f(r) = f_j, \qquad r_{j-1} < r \le r_j \quad (j = 1, 2, \ldots, N),$$

where $r_0 = R_1$ and $r_N = R_2$. The discretization then proceeds as follows. First, the integral of eq. (1a) is expressed as a summation by utilizing some appropriate quadrature approximation to give

$$\mathbf{A}\vec{f} = \vec{s}$$

where $\mathbf{A}$ is the $M \times N$ matrix of quadrature coefficients $w_j K(x_i, r_j)$ $(i = 1, 2, \ldots, M;\ j = 1, 2, \ldots, N)$, w_j are the weights associated with the quadrature rule, $\vec{f} = \left(f_1, f_2, \ldots, f_N\right)^T$ and $\vec{s} = \left(s(x_1), s(x_2), \ldots, s(x_M)\right)^T$. Second, the roughness measure $R_{\rm S}(f)$ is replaced by the following second difference approximation :

$$R = \sum_{j=2}^{N-1}\left(f_{j+1} - 2f_j + f_{j-1}\right)^2.$$

With these approximations, the smoothing functional of eq. (2a) can now be expressed as the following discrete representation:

$$\|\mathbf{W}\vec{d} - \mathbf{WA}\vec{f}\|_E^2 + \gamma \vec{f}^T \mathbf{H} \vec{f}, \tag{3}$$

where $\mathbf{W}$ is the $M \times M$ diagonal matrix

$$\mathbf{W} = \mathrm{diag}\{1/\sigma_1, 1/\sigma_2, \ldots, 1/\sigma_M\},$$

$\mathbf{H}$ is an $N \times N$ banded matrix of the form

$$\mathbf{H} = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 \\ -1 & 5 & -4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 \\ 1 & -4 & 6 & -4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 1 & -2 & 1 \end{pmatrix},$$

$\vec{d} = (D_1, D_2, \ldots, D_M)^T$ is the data vector and $\|\cdot\|_E$ denotes the usual Euclidean norm. Now, it is easily shown that the solution vector $\vec{f}$ which minimizes eq. (3) can be written as

$$\vec{f} = \left((\mathbf{WA})^T(\mathbf{WA}) + \gamma \mathbf{H}\right)^{-1} \left(\mathbf{WA}\right)^T \mathbf{W}\vec{d}. \tag{4}$$

13. The solution embodied in eq. (4) requires that an appropriate discretization be chosen for the PSDF $f(r)$. Indeed, the value of N must be chosen large enough so that there is no loss of fine structure in $f(r)$. On the other hand, although $f(r)$ can be represented to an arbitrary degree of accuracy by increasing the value of N, this incurs the disadvantage of increasing the dimension of the $(\mathbf{WA})^T(\mathbf{WA})$ matrix (cf. eq. (4)), thus making it more costly to invert on a computer. Hence, practical constraints imposed by the capabilities of the computer determine the upper bound on the value of N. Moreover, it is important to emphasize that certain kernel functions $K(x,r)$ such as the backscatter cross section are extremely oscillatory and erratic in form [14], with the result that a large value of N would be required in order to achieve a good quadrature approximation for $s(x)$ (cf. eq. (1a)). Under these circumstances, it is necessary for N to be large and, consequently, it is necessary to deal with the inversion of large dimensional matrices. In light of these problems, it would be desirable to obtain an analytical representation for the solution of eq. (2a) and thus do away with the need to implement a discrete approximation for $f(r)$. In so doing, discrete values of $f(r)$ will only be required for the quadrature approximation of the model optical data and not for the representation of the PSDF per se. The methodology for achieving this particular goal is presented in the next section.

ANALYTICAL REPRESENTATION OF THE SOLUTION FOR THE CONSTRAINED LINEAR INVERSION PROCEDURE

14. An explicit closed-form analytical representation for the form of the PSDF, which minimizes the smoothing functional $R_\gamma(f)$ of eq. (2a), can be obtained by application of certain standard techniques of Hilbert space theory [16]. From this perspective, it is convenient to visualize each admissible particle size distribution function $f(r)$ as an element of a real Hilbert space $\mathcal{H}$ consisting of all real-valued functions $h(r)$ defined on $\Omega \equiv [R_1, R_2]$ (i.e., the size domain) such that (1) $h' \equiv dh/dr$ is absolutely continuous and (2) $h'' \equiv d^2h/dr^2$ is an element of $\mathcal{L}^2(\Omega)$, the space of Lebesque square-integrable functions on Ω. It is convenient to equip the Hilbert space $\mathcal{H}$ with the following inner product:

$$\langle u, v \rangle = u(R_1)v(R_1) + u'(R_1)v'(R_1) + \int_\Omega u''(r)v''(r)\,dr,$$

where $u, v \in \mathcal{H}$. Note that the norm of an element in $\mathcal{H}$ is then given by $\|u\| = \langle u, u \rangle^{1/2}$. A brief comment is in order concerning the present choice of $\mathcal{H}$ and the associated inner product. In order to penalize undesirable undulations in $f(r)$, the constrained linear inversion procedure utilizes the integral of the second derivative of the PSDF over the size domain as the penalty functional (i.e., stabilizing functional). In view of this, $\mathcal{H}$ should be chosen as the class of all functions which possess a well-defined second derivative and the resulting norm (actually the norm inherited from the inner product) should be chosen so that it adequately reflects the "size" of the PSDF with regard to its "roughness". In this manner, the minimization in $\mathcal{H}$ using the given (i.e., chosen) norm then results in the penalization of the undesirable oscillatory behavior in $f(r)$ in an entirely natural fashion. With these definitions, the linear functional relationship between the theoretical datum $s(x_i)$ and the PSDF $f(r)$ (cf. eq. (1a)) can be viewed now as a continuous (bounded) linear functional L_i on $\mathcal{H}$ so, by virtue of the Riesz Representation Theorem [16], there must exist some $g_i \in \mathcal{H}$ such that

$$s(x_i) = L_i(f) \equiv \int_\Omega K(x_i, r)f(r)\,dr = \langle g_i, f \rangle \qquad (i = 1, 2, \ldots, M). \tag{5}$$

The elements g_i are called the representers for the bounded linear functionals L_i $(i = 1, 2, \ldots, M)$.

15. To find an analytic representation for the solution to eq. (2a), it is convenient to form two subspaces from $\mathcal{H}$. First, it is useful to form the finite-dimensional subspace spanned by the representers g_i $(i = 1, 2, \ldots, M)$ of the problem and define $\mathcal{G} = \mathrm{sp}\{g_1, g_2, \ldots, g_M\}$ where $\mathrm{sp}\{\cdot\}$ denotes the span of a set of elements in $\mathcal{H}$. Second, an inspection of the stabilizing functional $R_S(f)$ (cf. eq. (2c)) reveals that all elements of $\mathcal{H}$ which cause this functional to vanish (viz., $R_S(f) = 0$) must belong to a two-dimensional subspace of $\mathcal{H}$ spanned by the two elements

$$h_1(r) = 1 \qquad \text{and} \qquad h_2(r) = r - R_1.$$

Consequently, it is useful to define $\mathcal{K} = \mathrm{sp}\{h_1, h_2\}$. Note that elements from $\mathcal{K}$ are ideal particle size distribution functions in the sense that they are optimally smooth with respect to the stabilizing functional—viz., any element $k \in \mathcal{K}$ results in $R_S(k) = 0$ and, consequently, these elements are not penalized by the stabilizing functional.

16. Now, for a given value of the smoothing parameter γ, finding the function $f_\gamma(\mathbf{r}) \in \mathcal{H}$ which minimizes the smoothing functional $R_\gamma(f)$ is equivalent to finding the element of $\mathcal{H}$ which is "closest" to the subspace $\mathcal{K}$ of ideal functions and is, at the same time, consistent with the data misfit constraint embodied by $R_{\rm LS}(f)$. It is convenient to write the solution f_γ in the form

$$f_\gamma = \sum_{j=1}^{2} b_j h_j + \sum_{i=1}^{M} a_i g_i + \sigma, \tag{6}$$

where σ is an element of $\mathcal{H}$ which is orthogonal to both $\mathcal{G}$ and $\mathcal{K}$, viz. $\langle \sigma, g_i \rangle = 0$ $(i = 1, 2, \ldots, M)$ and $\langle \sigma, h_j \rangle = 0$ $(j = 1, 2)$. While the form for f_γ is quite intuitive, it should be noted that this form follows directly from the Decomposition Theorem [16] for Hilbert spaces. Note that the first term on the right hand side of eq. (6) is simply a representation of a general element from $\mathcal{K}$. Similarly, the second term on the right hand side of eq. (6) is a representation of a general element of $\mathcal{G}$. The element σ then represents that part of f_γ in $\mathcal{H}$ which does not lie in either $\mathcal{K}$ and/or $\mathcal{G}$. It will now be shown that σ must vanish if f_γ of eq. (6) is to minimize eq. (2a). Since f_γ is the "closest" element to $\mathcal{K}$ that is consistent with the data constraints, this requires that the length of the element $p \equiv \sum_{i=1}^{M} a_i g_i + \sigma$ be minimized subject to the data constraints (cf. eq. (6)). However, note that only the first term $\sum_{i=1}^{M} a_i g_i$ of p can affect the data misfit since the model data $s(x_i)$ are determined only by thc the g_i's (cf. eq. (5)) and $\langle \sigma, g_i \rangle = 0$ $(i = 1, 2, \ldots, M)$. Consequently, as far as the data misfit constraint is concerned, σ can be freely chosen. However, for minimum $||p||$, this requires that $\sigma = 0$. As a result, the PSDF $f_\gamma(\mathbf{r})$ which minimizes $R_\gamma(f)$ for a fixed value of the regularization parameter γ must possess the form

$$f_\gamma(\mathbf{r}) = \sum_{i=1}^{M} a_i g_i(\mathbf{r}) + \sum_{j=1}^{2} b_j h_j(\mathbf{r}). \tag{7}$$

17. The undetermined coefficients a_i $(i = 1, 2, \ldots, M)$ and b_j $(j = 1, 2)$ in the representation of eq. (7) can be determined as follows. If eq. (7) is substituted into eqs. (2a), (2b), and (2c), it is relatively straightforward to show that the smoothing functional $R_\gamma(f)$ can be expressed as the quadratic form

$$R_\gamma(f) = M^{-1}\left(\vec{d} - \mathbf{\Gamma}\vec{a} - \mathbf{S}\vec{b}\right)^T \mathbf{W}^2 \left(\vec{d} - \mathbf{\Gamma}\vec{a} - \mathbf{S}\vec{b}\right) + \gamma \vec{a}^T \mathbf{\Gamma} \vec{a}, \tag{8a}$$

where $\mathbf{\Gamma}$ is the $M \times M$ Gram matrix defined as

$$\mathbf{\Gamma} = \left(\Gamma_{ij}\right)_{i=1,2,\ldots,M}^{j=1,2,\ldots,M}, \qquad \Gamma_{ij} \equiv \langle g_i, g_j \rangle, \tag{8b}$$

$\mathbf{S}$ is the $M \times 2$ matrix defined as

$$\mathbf{S} = \left(S_{ij}\right)_{i=1,2,\ldots,M}^{j=1,2}, \qquad S_{ij} \equiv \langle g_i, h_j \rangle, \tag{8c}$$

and $\vec{a}$ and $\vec{b}$ are M- and 2-tuples, respectively, given by

$$\vec{a} = \left(a_1, a_2, \ldots, a_M\right)^T \qquad \text{and} \qquad \vec{b} = \left(b_1, b_2\right)^T. \tag{8d}$$

The data vector $\vec{d}$ and the diagonal matrix $\mathbf{W}$ have already been defined with reference to eq. (3). Now, differentiating $R_\gamma(f)$ in eq. (8a) with respect to the coefficient or parameter vectors $\vec{a}$ and $\vec{b}$ and setting the results to zero, yields the following system of linear equations:

$$(\mathbf{\Gamma} + M\gamma\mathbf{W}^{-2})\vec{a} + \mathbf{S}\vec{b} = \vec{d}, \tag{9a}$$

$$\mathbf{S}^T\vec{a} = \vec{0}; \tag{9b}$$

or, equivalently, in matrix form

$$\begin{pmatrix} \mathbf{\Gamma} + M\gamma\mathbf{W}^{-2} & \mathbf{S} \\ \mathbf{S}^T & 0 \end{pmatrix} \begin{pmatrix} \vec{a} \\ \vec{v} \end{pmatrix} = \begin{pmatrix} \vec{d} \\ \vec{0} \end{pmatrix}, \tag{9c}$$

where 0 is a 2×2 zero matrix and $\vec{0}$ is a 2×1 zero vector. The solution of eq. (9c) provides the values of the unknown coefficients $\{a_i\}_{i=1}^{M}$ and $\{b_j\}_{j=1}^{2}$.

18. Explicit expressions for the coefficient vectors $\vec{a}$ and $\vec{b}$ can be derived as follows. First, note that the matrix $(\mathbf{\Gamma} + M\gamma\mathbf{W}^{-2})$ is nonsingular since (1) the Gram matrix $\mathbf{\Gamma}$ is symmetric and non-negative definite by definition (cf. eq. (8b)), (2) $\mathbf{W}$ is positive definite, and (3) γ is positive by assumption (i.e., recall that γ is the positive regularization parameter). Consequently, if eq. (9a) is premultiplied by $\mathbf{S}^T(\mathbf{\Gamma} + M\gamma\mathbf{W}^{-2})^{-1}$, one obtains

$$\mathbf{S}^T\vec{a} + \mathbf{S}^T(\mathbf{\Gamma} + M\gamma\mathbf{W}^{-2})^{-1}\mathbf{S}\vec{b} = \mathbf{S}^T(\mathbf{\Gamma} + M\gamma\mathbf{W}^{-2})^{-1}\vec{d},$$

which can be expressed as

$$\vec{b} = \left[\mathbf{S}^T(\mathbf{\Gamma} + M\gamma\mathbf{W}^{-2})^{-1}\mathbf{S}\right]^{-1}\mathbf{S}^T(\mathbf{\Gamma} + M\gamma\mathbf{W}^{-2})^{-1}\vec{d}, \tag{10a}$$

on using eq. (9b) and assuming that the matrix $\mathbf{S}$ has full rank. Now, $\vec{a}$ can be computed from

$$\vec{a} = (\mathbf{\Gamma} + M\gamma\mathbf{W}^{-2})^{-1}(\vec{d} - \mathbf{S}\vec{b}) \tag{10b}$$

by virtue of eqs. (9a) and (10a). Eqs. (10a) and (10b) provide the coefficients of the desired solution to the constrained linear inversion problem. Indeed, eqs. (10a) and (10b) in conjunction with eq. (7) provide a continuous representation for the PSDF $f_\gamma(r)$ which minimizes $R_\gamma(f)$. However, to be able to utilize this continuous representation, it is necessary to develop explicit expressions for the representers g_i $(i = 1, 2, \ldots, M)$. The solution of this problem is the subject of the next section.

EXPLICIT COMPUTATION OF THE REPRESENTERS FOR THE CONSTRAINED LINEAR INVERSION PROCEDURE

19. Although the Riesz Representation Theorem guarantees that every bounded linear functional L_i defined on a Hilbert space $\mathcal{H}$ possesses a representer $g_i \in \mathcal{H}$ such that $L_i(f) = \langle g_i, f \rangle$ for every $f \in \mathcal{H}$, the

theorem does not state how such a function can be constructed in general. Hence, it is useful to approach the problem of the construction of the representers for the constrained linear inversion procedure in the following heuristic manner. To this end, it is convenient to view the Hilbert space $\mathcal{H}$ (defined in the previous section) as the direct sum of two Hilbert spaces $\mathcal{H}_0$ and $\mathcal{H}_1$ (viz., $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$) defined as follows:

$$\mathcal{H}_0 = \left\{ f : df/dr \text{ is absolutely continuous, } d^2f/dr^2 = 0 \right\}$$

and

$$\mathcal{H}_1 = \left\{ f : df/dr \text{ is absolutely continuous, } d^2f/dr^2 \in \mathcal{L}^2(\Omega), f(R_1) = 0, df(R_1)/dr = 0 \right\}.$$

Observe that $\mathcal{H}_0$ is a two-dimensional subspace of $\mathcal{H}$ which coincides with the subspace $\mathcal{K}$ defined earlier. The inner products on $\mathcal{H}_0$ and $\mathcal{H}_1$ must be defined as

$$\langle f, g \rangle_0 \equiv f(R_1)g(R_1) + f'(R_1)g'(R_1), \qquad f, g \in \mathcal{H}_0$$

and

$$\langle f, g \rangle_1 \equiv \int_\Omega \frac{d^2f(r)}{dr^2} \frac{d^2g(r)}{dr^2}\, dr, \qquad f, g \in \mathcal{H}_1,$$

respectively, so as to be consistent with the inner product chosen for $\mathcal{H}$ (cf. previous section for the definition of this inner product). Then, $\langle f, g \rangle = \langle P_0 f, P_0 g \rangle_0 + \langle P_1 f, P_1 g \rangle_1$ $(f, g \in \mathcal{H})$, where P_0 and P_1 denote the projection operators from $\mathcal{H}$ onto $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively.

20. As the first step in the construction of the representers g_i, consider the problem of finding functions $k_{r'}^{(0)} \in \mathcal{H}_0$ and $k_{r'}^{(1)} \in \mathcal{H}_1$ such that

$$f(r') = \langle k_{r'}^{(0)}, f \rangle_0, \qquad \text{for any } f \in \mathcal{H}_0 \tag{11a}$$

and

$$f(r') = \langle k_{r'}^{(1)}, f \rangle_1, \qquad \text{for any } f \in \mathcal{H}_1. \tag{11b}$$

In the preceding equations, r' is an arbitrary fixed point in Ω. It should be noted that $k_{r'}^{(0)}$ and $k_{r'}^{(1)}$ are simply the representers for the point evaluation functionals in $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. Due to the simplicity of the inner product on $\mathcal{H}_0$ and in view of the fact that $\mathcal{H}_0$ is a two-dimensional space spanned by the functions $h_1(r) \equiv 1$ and $h_2(r) \equiv (r - R_1)$, a little reflection shows that the function $k_{r'}^{(0)}$ possesses the form

$$k_{r'}^{(0)}(r) = 1 + (r - R_1)(r' - R_1), \qquad r, r' \in \Omega. \tag{12}$$

That this form for $k_{r'}^{(0)}(r)$ is correct can be readily verifed by substitution into eq. (11a).

21. The function $k_{r'}^{(1)}(r)$ is more difficult to determine, but can be found as follows. First, note that eq. (11b) can be expressed as

$$f(r') = \int_\Omega \frac{d^2k_{r'}^{(1)}(r)}{dr^2} \frac{d^2f(r)}{dr^2}\, dr \equiv \int_\Omega G(r', r) \frac{d^2f(r)}{dr^2}\, dr, \tag{13a}$$

where $G(r', r)$ is a Green's function determined such that

$$\frac{d^2G(r', r)}{dr^2} = \delta(r - r'), \qquad G(r', R_1) = 0 = \frac{dG(r', R_1)}{dr}, \tag{13b}$$

where $\delta(\cdot)$ denotes the Dirac delta function. The solution of the preceding boundary value problem (i.e., eq. (13b)) can be easily shown to be given by

$$G(r', r) = \begin{cases} 0, & \text{if } R_1 \le r \le r'; \\ r - r', & \text{if } R_2 \ge r \ge r'. \end{cases} \tag{14}$$

That eq. (14) is valid can be easily verified by substituting this expression into eq. (13a) and integrating the result by parts twice. Now, in light of eq. (13a), $k_{r'}^{(1)}(r)$ is related to the Green's function as

$$\frac{d^2k_{r'}^{(1)}(r)}{dr^2} = G(r', r)$$

or, equivalently, as

$$k_{r'}^{(1)}(r) = \int_\Omega G(r', r'')G(r, r'')\, dr''. \tag{15}$$

Inserting eq. (14) into eq. (15) and evaluating the result yields

$$k_{r'}^{(1)}(r) = \begin{cases} \displaystyle\int_{R_1}^{r} (r' - r'')(r - r'')\, dr'' = (r - R_1)^2(r' - R_1)/2 - (r - R_1)^3/6, & \text{if } R_1 \le r \le r'; \\ \displaystyle\int_{R_1}^{r'} (r' - r'')(r - r'')\, dr'' = (r - R_1)(r' - R_1)^2/2 - (r' - R_1)^3/6, & \text{if } R_2 \ge r \ge r'. \end{cases} \tag{16}$$

22. Having determined the representers for the point evaluation functionals in $\mathcal{H}_0$ and $\mathcal{H}_1$, it is now a simple matter to compute the representers for an arbitrary bounded linear functional in $\mathcal{H}$. In particular, one is interested in the representers g_i associated with the data functionals L_i defined in eq. (5) ($i = 1, 2, \ldots, M$). To this purpose, first note that by virtue of eqs. (11a) and (11b), $f(r') = \langle k_{r'}^{(0)}, P_0 f\rangle_0 + \langle k_{r'}^{(1)}, P_1 f\rangle_1 \equiv \langle k_{r'}, f\rangle$ with $k_{r'} \equiv k_{r'}^{(0)} + k_{r'}^{(1)}$ ($f \in \mathcal{H}$). Consequently,

$$s(x_i) = \langle g_i, f\rangle \equiv L_{i(r')}\big(f(r')\big) = L_{i(r')}\langle k_{r'}, f\rangle = \langle L_{i(r')}k_{r'}, f\rangle,$$

so

$$g_i(r) = L_{i(r')}\big(k_{r'}(r)\big) = \int_\Omega K(x_i, r')k_{r'}(r)\, dr', \tag{17a}$$

where

$$k_{r'}(r) = k_{r'}^{(0)}(r) + k_{r'}^{(1)}(r) \tag{17b}$$

and $L_{i(r')}(\cdot)$ denotes the i-th data functional (cf. eq. (5)) operating on what follows (viz., $(\cdot)$) considered as a function of r'. Now, substitution of the functional forms for $k_{r'}^{(0)}$ and $k_{r'}^{(1)}$ exhibited by eqs. (12) and (16), respectively, into eq. (17a) leads to the following explicit expression for the representers g_i:

$$\begin{aligned} g_i(r) = {} & \int_\Omega K(x_i, r')\, dr' + (r - R_1)\int_\Omega (r' - R_1)K(x_i, r')\, dr' \\ & + \frac{1}{2}(r - R_1)\int_{R_1}^{r} (r' - R_1)^2 K(x_i, r')\, dr' - \frac{1}{6}\int_{R_1}^{r} (r' - R_1)^3 K(x_i, r')\, dr' \\ & + \frac{1}{2}(r - R_1)^2\int_{r}^{R_2} (r' - R_1)K(x_i, r')\, dr' - \frac{1}{6}(r - R_1)^3\int_{r}^{R_2} K(x_i, r')\, dr'. \end{aligned} \tag{18}$$

23. With the analytical form for the representers g_i given in eq. (18), it is now possible to compute the analytical solution for the constrained linear inversion procedure using eqs. (7), (8b), (8c), (10a), and (10b). To simplify the computational procedure somewhat, it is important to observe that the first two terms of the expression for $g_i(r)$ given by eq. (18) are elements from $\mathcal{K}$ and, consequently, can be incorporated with the term $\sum_{j=1}^{2} b_j h_j(r) \in \mathcal{K}$ in the representation of $f_\gamma(r)$. Hence, it is convenient to write the representation for $f_\gamma(r)$ as

$$f_\gamma(r) = \sum_{i=1}^{M} a_i \bar{g}_i(r) + \sum_{j=1}^{2} b_j h_j(r), \tag{19a}$$

where

$$\begin{aligned}\bar{g}_i(r) = {} & \frac{1}{2}(r - R_1)\int_{R_1}^{r} (r' - R_1)^2 K(x_i, r')\,dr' - \frac{1}{6}\int_{R_1}^{r} (r' - R_1)^3 K(x_i, r')\,dr' \\ & + \frac{1}{2}(r - R_1)^2 \int_{r}^{R_2} (r' - R_1) K(x_i, r')\,dr' - \frac{1}{6}(r - R_1)^3 \int_{r}^{R_2} K(x_i, r')\,dr'.\end{aligned} \tag{19b}$$

Observe that $\bar{g}_i(r)$, given by eq. (19b), is simply $g_i(r)$ with the component in $\mathcal{K}$ removed (viz., $\bar{g}_i = P_1 g_i$). The coefficient vectors $\vec{a}$ and $\vec{b}$, which determine $f_\gamma(r)$ in eq. (19a), are still given by eqs. (10a) and (10b). However, in this case, the elements of the Gram matrix Γ are given now by

$$\langle g_i, \bar{g}_j \rangle = \langle \bar{g}_i, \bar{g}_j \rangle = \int_\Omega \frac{d^2 \bar{g}_i(r)}{dr^2} \frac{d^2 \bar{g}_j(r)}{dr^2}\,dr, \qquad i, j = 1, 2, \ldots, M, \tag{20a}$$

with

$$\frac{d^2 \bar{g}_i(r)}{dr^2} = \int_r^{R_2} (r' - R_1) K(x_i, r')\,dr' - (r - R_1)\int_r^{R_2} K(x_i, r')\,dr'. \tag{20b}$$

Furthermore, the elements of the S matrix are unchanged and are given by

$$\langle g_i, h_j \rangle = L_i(h_j) = \int_\Omega K(x_i, r) h_j(r)\,dr, \qquad i = 1, 2, \ldots, M; j = 1, 2. \tag{20c}$$

24. In summary then, for a given value of the regularization parameter γ, the solution of the constrained linear inversion procedure (i.e., the particle size distribution function which minimizes the smoothing functional of eq. (2a)) possesses the continuous representation

$$f_\gamma(r) = \vec{a}^T \vec{\bar{g}}(r) + \vec{b}^T \vec{h}(r), \qquad r \in \Omega, \tag{21a}$$

where $\vec{\bar{g}}(r)$ and $\vec{h}(r)$ are vector-valued functions on Ω defined as

$$\vec{\bar{g}}(r) \equiv \left(\bar{g}_1(r), \bar{g}_2(r), \ldots, \bar{g}_M(r)\right)^T \tag{21b}$$

and

$$\vec{h}(r) \equiv \left(h_1(r), h_2(r)\right)^T. \tag{21c}$$

The coefficient or parameter vectors $\vec{a}$ and $\vec{b}$ of eq. (21a) are computed according to eqs. (10a) and (10b) with the matrix elements of Γ and S determined as per eqs. (20a), (20b), and (20c). This solution for the

PSDF (cf. eqs. (21a), (21b), and (21c)) should be contrasted with the conventional approximate form of the solution represented by eq. (4). Eq. (4) is based on a discrete approximation (i.e., a piece-wise constant approximation) of the PSDF which requires the selection of an appropriate number of discrete size intervals (N) to use for the representation of the continuous aerosol particle size distribution. On the other hand, eqs. (21a)–(21c) constitute a continuous representation of the PSDF which solves the constrained linear inversion problem. With this solution, there is no direct need to utilize a discrete approximation for $f(r)$ with its concomitant subjectivity in the selection of N. It is important to note that closed-form expressions cannot be obtained for the model optical data $s(x_i)$ (cf. eq. (1a)) or for the matrix elements Γ_{ij} of Γ (cf. eqs. (20a) and (20b)) and S_{ij} of S (cf. eq. (20c)), so that some form of quadrature rule is required in the computation. However, even though the quadrature approximation for $s(x_i)$ must necessarily involve the discretization of the PSDF $f(r)$, it should be emphasized that this discretization is only utilized to obtain an approximation for the model optical data and is not related to the representation of $f(r)$ per se. In the usual implementation of the constrained linear inversion procedure, the representation of $f(r)$ as a piecewise-constant function is directly associated with the quadrature approximation for $s(x_i)$. In the present implementation of this procedure, a continuous and exact representation is provided for $f(r)$ and a discretization of this representation is utilized only for the quadrature approximation of $s(x_i)$—viz., the discretization is not utilized for the approximate representation of $f(r)$ as in the usual implementation. Finally, the computation of the PSDF via eq. (21a) involves the inversion of a matrix of order M, whereas the computation of the PSDF via eq. (4) requires the inversion of a matrix of order N. In the problem of the inference of the PSDF from light scattering data, N (number of size intervals) is almost always larger than M (number of data points).

SELECTION OF THE REGULARIZATION PARAMETER

25. Before a complete solution can be given for the constrained linear inversion procedure, it is necessary to address the problem of the proper selection of the regularization (smoothing) parameter γ. In this paper, it is proposed that the generalized cross-validation (GCV) method be utilized for the selection of this parameter. The GCV method is a rotation-invariant form of the ordinary cross-validation method introduced by Allen [17]. Indeed, the GCV method has been successfully applied to determining the correct degree of smoothing of discrete, noisy data using spline functions [18,19] and to the optimal smoothing of probability density [20] and spectral density estimates [21]. For a detailed theoretical treatment of the properties of the GCV estimate of a smoothing parameter, the reader should consult Craven and Wahba [19].

26. With reference to the problem treated in this paper, the GCV method specifies that the optimal estimate $\hat{\gamma}$ of the smoothing parameter γ be chosen to minimize the criterion function

$$V(\gamma) = \frac{1}{M} \sum_{k=1}^{M} \frac{\left(D_k - L_k\left(f_\gamma^{(k)}\right)\right)^2}{\sigma_k^2} w_k^\gamma, \tag{22a}$$

where $f_\gamma^{(k)}$ denotes the minimizer of the smoothing functional $R_\gamma(f)$ (cf. eq. (2a)) with the k-th data point (i.e., D_k) omitted from the computation. Recall that L_k ($k = 1, 2, \ldots, M$) are the data functionals defined as per eq. (5). Moreover, in eq. (22a), $w_k^{(\gamma)}$ are weights given by

$$w_k^{(\gamma)} = \Big(1 - p_{kk}(\gamma)\Big) \Big/ \Big(1 - \frac{1}{M}\mathrm{tr}\Big(\mathbf{P}(\gamma)\Big)\Big), \tag{22b}$$

where $p_{kk}(\gamma)$ is the kk-th entry of the $M \times M$ matrix $\mathbf{P}(\gamma)$ defined as

$$\begin{pmatrix} L_1(f_\gamma) \\ L_2(f_\gamma) \\ \vdots \\ L_M(f_\gamma) \end{pmatrix} = \mathbf{P}(\gamma)\vec{d}, \tag{23}$$

and $\mathrm{tr}(\cdot)$ denotes the trace operation. It should be noted that the basic logic behind the selection of γ to minimize $V(\gamma)$ stems from the fact that this choice for γ minimizes, on the average, the squared error between the data points D_k ($k = 1, 2, \ldots, M$) and the predictors $L_k(f_\gamma^{(k)})$ of these points. Observe that the predictor for the k-th data point is obtained using the estimate $f_\gamma^{(k)}$ of the PSDF which does not involve this particular data point (i.e., $f_\gamma^{(k)}$ is the solution to the constrained linear inversion procedure with the k-th data point omitted).

27. The form of $V(\gamma)$ given in eqs. (22a) and (22b) is somewhat complex and is not suited for numerical computations. However, it can be shown [19] that $V(\gamma)$ can be expressed in the following form that is more amenable to computations:

$$V(\gamma) = \frac{M^{-1}\|\mathbf{W}(\mathbf{I} - \mathbf{P}(\gamma))\vec{d}\|_E^2}{\Big[M^{-1}\mathrm{tr}\big(\mathbf{I} - \mathbf{P}(\gamma)\big)\Big]^2}. \tag{24}$$

It should be emphasized that $\|\cdot\|_E$ in eq. (24) denotes the usual Euclidean norm and not the norm $\|\cdot\|$ inherited from the inner product on $\mathcal{H}$. The remaining problem now is to determine the form of the matrix $\mathbf{P}(\gamma)$ for the present problem and to substitute it into eq. (24). To this end, first note that (cf. eq. (19a))

$$\begin{aligned} L_k(f_\gamma(r)) &= \sum_{i=1}^{M} a_i L_k(\bar{g}_i(r)) + \sum_{j=1}^{2} b_j L_k(h_j(r)) \\ &= \sum_{i=1}^{M} a_i \langle \bar{g}_k, \bar{g}_i \rangle + \sum_{j=1}^{2} b_j \langle g_k, h_j \rangle. \end{aligned}$$

In view of the preceding expression and the definition of the matrix $\mathbf{P}(\gamma)$ given in eq. (23), it follows that

$$\begin{pmatrix} L_1(f_\gamma) \\ L_2(f_\gamma) \\ \vdots \\ L_M(f_\gamma) \end{pmatrix} = \mathbf{\Gamma}\vec{a} + \mathbf{S}\vec{b} = \mathbf{P}(\gamma)\vec{d},$$

so

$$\Big(\mathbf{I} - \mathbf{P}(\gamma)\Big)\vec{d} = M\gamma\mathbf{W}^{-2}\vec{a},$$

on using eq. (9a). In view of this equation, it follows that the numerator of $V(\gamma)$, as displayed in eq. (24), can be expressed as follows:

$$\|\mathbf{W}(\mathbf{I}-\mathbf{P}(\gamma))\vec{d}\|_E^2 = (M\gamma)^2\vec{a}^T\mathbf{W}^{-2}\vec{a}. \tag{25}$$

28. To proceed further, it would be necessary to substitute the expression for $\vec{a}$ displayed in eq. (10b) into eq. (25) and then try to manipulate the result into a computationally useful form. However, it turns out that the expression of $\vec{a}$, as embodied in eq. (10b), is not the most useful form for this purpose. As a result, it is sensible to consider an alternative, but equivalent, expression for $\vec{a}$ that will prove to be more amenable to simplifications, at least as far as the evaluation of eq. (25) is concerned. To this end, first observe from eq. (9b) that $\vec{a}$ must lie in the null (or annihilator) space of $\mathbf{S}^T$ so it is possible to express $\vec{a}$ as

$$\vec{a} = \mathbf{Q}\vec{v} \qquad \text{where} \qquad \mathbf{S}^T\mathbf{Q} = \mathbf{0} \quad \text{and} \quad \vec{v} \in R^{M-2}. \tag{26}$$

In eq. (26), $\mathbf{Q}$ is an $M \times (M-2)$ matrix whose columns span the null space of $\mathbf{S}^T$. Again, as in the derivation of eq. (10a), it is assumed that the rank of $\mathbf{S}$ is two (i.e., $\mathbf{S}$ is full rank). Now, if eq. (9a) is premultiplied by $\mathbf{Q}^T$ and combined with eq. (26), it is easily shown that

$$\vec{a} = \mathbf{Q}\left[\mathbf{Q}^T\left(\mathbf{\Gamma} + M\gamma\mathbf{W}^{-2}\right)\mathbf{Q}\right]^{-1}\mathbf{Q}^T\vec{d}. \tag{27a}$$

Moreover, if eq. (9a) is premultiplied by $\mathbf{S}^T\mathbf{W}^2$ and combined with eq. (9b), the following expression for $\vec{b}$ results:

$$\vec{b} = \left(\mathbf{S}^T\mathbf{W}^2\mathbf{S}\right)^{-1}\mathbf{S}^T\mathbf{W}^2\left(\vec{d} - \mathbf{\Gamma}\vec{a}\right). \tag{27b}$$

It should be noted that eqs. (26), (27a) and (27b) constitute alternative, but equivalent, expressions for $\vec{a}$ and $\vec{b}$ contained in eqs. (10a) and (10b).

29. It is important to point out that the matrix $\mathbf{Q}$ is not uniquely determined by eq. (26). Consequently, it is possible to utilize this fact to choose a convenient form for $\mathbf{Q}$ that permits a simplification in the evaluation of $\|\mathbf{W}(\mathbf{I}-\mathbf{P}(\gamma))\vec{d}\|_E^2$. In this connection, it is convenient to choose $\mathbf{Q}$ so that its columns span the null space of $\mathbf{S}^T$ (cf. eq. (26)) and satisfies

$$\mathbf{Q}^T\mathbf{W}^{-2}\mathbf{Q} = \mathbf{I}. \tag{28a}$$

With this choice, insertion of eq. (27a) into eq. (25) followed by some simple matrix manipulations leads to the following expression:

$$\begin{aligned}\|\mathbf{W}(\mathbf{I}-\mathbf{P}(\gamma))\vec{d}\|_E^2 &= (M\gamma)^2\vec{\xi}^T(\mathbf{\Lambda} + M\gamma\mathbf{I})^{-2}\vec{\xi} \\ &= (M\gamma)^2\sum_{i=1}^{M-2}\frac{\xi_i}{(\lambda_i + M\gamma)^2}.\end{aligned} \tag{28b}$$

In eq. (28a), an eigenvector-eigenvalue decomposition has been introduced for $\mathbf{Q}^T\mathbf{\Gamma}\mathbf{Q}$, viz.

$$\mathbf{Q}^T\mathbf{\Gamma}\mathbf{Q} \equiv \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T, \tag{28c}$$

where $\mathbf{U}$ is an $(M-2)\times(M-2)$ orthogonal matrix whose columns are composed of the eigenvectors of $\mathbf{Q}^T\boldsymbol{\Gamma}\mathbf{Q}$ and

$$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{M-2})^T, \tag{28d}$$

is an $(M-2)\times(M-2)$ diagonal matrix whose diagonal elements are composed of the eigenvalues λ_i $(i = 1, 2, \ldots, (M-2))$ of $\mathbf{Q}^T\boldsymbol{\Gamma}\mathbf{Q}$. Moreover, in eq. (28b), the vector $\vec{\xi}$ is defined as

$$\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_{M-2})^T \equiv \mathbf{U}^T\mathbf{Q}^T\vec{d}. \tag{28e}$$

In a similar manner, it is possible to show that the denominator of eq. (24) can be written as follows:

$$\begin{aligned} \mathrm{tr}(\mathbf{I} - \mathbf{P}(\gamma)) &= (M\gamma)\mathrm{tr}((\boldsymbol{\Lambda} + M\gamma\mathbf{I})^{-1}) \\ &= (M\gamma)\sum_{i=1}^{M-2}\frac{1}{\lambda_i + M\gamma}. \end{aligned} \tag{29}$$

Now substitution of eqs. (28b) and (29) into eq. (24) finally leads to the following expression for $V(\gamma)$:

$$V(\gamma) = \frac{M\sum_{i=1}^{M-2}\left(\xi_i/(\lambda_i + M\gamma)\right)^2}{\left[\sum_{i=1}^{M-2} 1/(\lambda_i + M\gamma)\right]^2}. \tag{30}$$

In light of eq. (30), it is now relatively easy to obtain a value of γ that minimizes $V(\gamma)$ by using a golden section search in one dimension. This optimal value for γ can then be substituted into eqs. (10a) and (10b), or equivalently, into eqs. (27a) and (27b) for the computation of $\vec{a}$ and $\vec{b}$. Once this is completed, the solution of the constrained linear inversion procedure can then be obtained by evaluation of eq. (21a).

ALGORITHM FOR THE AUTOMATIC RECOVERY OF THE PSDF

30. In view of the results obtained in the previous sections, the following algorithm is proposed for the automatic reconstruction of the PSDF, given a finite number M of discrete, noisy light scattering data D_i along with the standard deviations σ_i in their measurement. A flowchart of the implementation used to realize the automatic recovery of the PSDF is given in Fig. 1.

31. *Step 1:* Compute the matrix elements of $\boldsymbol{\Gamma}$ and $\mathbf{S}$ as per eqs. (20a), (20b), and (20c). Since the Mie scattering kernels are complicated functions, these matrix elements cannot be evaluated analytically. Rather, the computation of these matrix elements must be effected by utilizing an appropriate numerical quadrature rule.

32. *Step 2:* Compute the QR decomposition [22] of $\mathbf{WS}$, viz.

$$\mathbf{WS} = \begin{pmatrix} \mathbf{Q}_1 & \vdots & \mathbf{Q}_2 \end{pmatrix}\begin{pmatrix} \mathbf{R}_1 \\ \cdots \\ \mathbf{0} \end{pmatrix},$$

where $\mathbf{R}_1$ is a 2×2 upper triangular matrix and $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are $M\times 2$ and $M\times(M-2)$ matrices, respectively, such that $\mathbf{Q}_1^T\mathbf{Q}_1 = \mathbf{I}$, $\mathbf{Q}_1^T\mathbf{Q}_2 = \mathbf{0}$, and $\mathbf{Q}_2^T\mathbf{Q}_2 = \mathbf{I}$. Now observe that matrix $\mathbf{WQ}_2$ can be identified with $\mathbf{Q}$ in eqs. (26) and (28a) since $\mathbf{Q}_2^T\mathbf{WS} = \mathbf{0}$ (or, equivalently, $\mathbf{S}^T(\mathbf{WQ}_2) = \mathbf{0}$) and $\mathbf{Q}_2^T\mathbf{Q}_2 \equiv \mathbf{Q}^T\mathbf{W}^{-2}\mathbf{Q} = \mathbf{I}$.

33. *Step 3:* Compute the eigenvector-eigenvalue decomposition of $\mathbf{Q}^T\mathbf{\Gamma Q}$ (cf. eq. (28c)), construct the vector $\vec{\xi}$ as per eq. (28e), and compute the GCV criterion function $V(\gamma)$ using eq. (30). Now apply a golden section search [23] in one dimension to determine the non-negative value of the smoothing parameter γ that minimizes $V(\gamma)$.

34. *Step 4:* Using the value of γ determined in Step 3, compute the parameter vectors $\vec{a}$ and $\vec{b}$ using either eqs. (27a) and (27b) or eqs. (10a) and (10b). However, since $\mathbf{Q}$ has already been computed in Step 2, it is perhaps more convenient to use eqs. (27a) and (27b) for this computation. Consequently, eqs. (27a) and (27b) are used in the computation for the present implementation.

35. *Step 5:* Evaluate the PSDF $f_\gamma(r)$ (i.e., the solution of the constrained linear inversion procedure) using eq. (21a).

36. It is important to note that the bulk of the computations in the preceding algorithm is contained in Step 1. However, once an experimental design has been specified (i.e., once the wavelengths and scattering angles at which the scattered light is to be measured have been selected), it is only necessary to compute the matrix elements of $\mathbf{\Gamma}$ and $\mathbf{S}$ once since these elements are independent of the data. These results can then be saved for utilization in the inversion of all subsequent data sets that adhere to the given experimental arrangement. Generally speaking, the incorporation of the present algorithm into an Mie scattering-based particle sizing instrument would involve the a priori computation of the elements of $\mathbf{\Gamma}$ and $\mathbf{S}$ based on the experimental configuration of the instrument. These matrix elements would then be stored in the onboard computer and the implementation of the PSDF inversion algorithm would then essentially involve the computation of Steps 2 to 5. Moreover, the computation of the matrix elements in Step 1 and of $f_\gamma(r)$ in Step 5 can be efficiently implemented using lookup tables for the relevant Mie kernels. This would allow for a rapid real-time PSDF inversion using only the capabilities of modern desktop computers.

SOME NUMERICAL EXAMPLES

37. If $s(x)$ in eq. (1a) is identified with the optical extinction of light at the wavelength $x \equiv \lambda$ for a homogenous, polydispersed aerosol cloud characterized by the particle size distribution function $f(r)$ (or, equivalently, the normalized number concentration of aerosol particles in the size interval r to $r + dr$), then the kernel function $K(x,r)$, corresponding to the Fredholm integral operator that maps $f(r)$ to $s(x)$, is given by

$$K(x,r) = \pi r^2 Q_{\text{ext}}(x,r), \tag{31a}$$

where $Q_{ext}(x,r)$ is the extinction efficiency factor evaluated using Mie theory [13,14] for a particle of radius r and wavelength x. Furthermore, if $s(x)$ is identified with the light scattering intensity (i.e., radiance) at the scattering angle $x \equiv \theta$ obtained for a fixed wavelength λ_0, then the kernel function $K(x,r)$ that transforms $f(r)$ to $s(x)$ via the Fredholm integral operator assumes the form

$$K(x,r) = \left(\frac{\lambda_0}{2\pi}\right)^2 \left(\frac{i_1(x,r) + i_2(x,r)}{2}\right), \tag{31b}$$

where $i_1(x,r)$ and $i_2(x,r)$ are the Mie intensity parameters [13,14] for scattered light with perpendicular and parallel polarization, respectively. It should be noted that the extinction efficiency factor as well as the Mie intensity parameters are complicated functions of r and x which are usually expressed as infinite sums of terms involving the Ricatti-Bessel functions and the Legendre polynomials. Nevertheless, there now exist very efficient and robust procedures for the evaluation of the Mie kernels [14,24]. Strictly speaking, it is important to note that the extinction and scattering parameters also depend on the complex index of refraction m of the particles. In the following examples, the index of refraction of the particles is assumed to be known a priori. In particular, all simulations considered are performed assuming that the particles have a real (viz., the particles are non-absorbing) refractive index $m = 1.54$.

38. Optical extinction (or turbidity) data were calculated at 15 discrete wavelengths spanning the interval between $x = 0.1\ \mu$m and $x = 7.5\ \mu$m for various functional forms of the PSDF $f(r)$. The radii limits which determine the extent of the size domain for $f(r)$ are taken to be $\Omega \equiv [R_1, R_2] = [0.1\,\mu\text{m}, 5.0\,\mu\text{m}]$. In addition, light scattering intensity data were generated at $\lambda_0 = 0.65\ \mu$m for 15 scattering angles between 0 and 90 degrees. Since eq. (1a) with the kernel functions of eqs. (31a) and (31b) cannot be integrated analytically due to the complexity of the Mie kernels, it is necessary to calculate the optical data numerically by utilizing some quadrature rule. The present simulations were generated using Simpson's rule as the quadrature formula. These simulated optical data were then degraded with pseudorandom normally distributed noise adjusted to correspond to a prescribed level of root-mean-square (RMS) error. These degraded spectral extinction and scattered radiance data comprised the components of the data vector $\vec{d}$ (cf. eq. (1b)) of dimension $M = 30$ that served as the input to the inversion algorithm.

39. Fig 2 displays the result of the reconstruction of a uniform PSDF given by $f(r) = 1.0$ for $r \in \Omega$. The figure illustrates that for the case of noise-free multispectral extinction and scattered radiance data, the recovered PSDF coincides exactly with the true (actual) PSDF. The regularization (smoothing) parameter, computed using the generalized cross-validation procedure, was effectively zero (to within the roundoff errors of the computer) for the case of noise-free data. Moreover, an examination of the coefficient (parameter) vectors returned by the retrieval algorithm revealed that $\vec{a} = \vec{0}$ and $\vec{b} = (1.0, 0.0)^T$, a result which would yield $f(r) = 1.0$ exactly as per eq. (21a). The inversion result of this particular example is interesting because it illustrates the fact that if a solution can be constructed from elements in $\mathcal{K}$ that satisfy the data constraints, that solution is preferably chosen by the algorithm. Recall that the elements in $\mathcal{K}$ are the

preferred solution candidates since these elements are considered to be optimally smooth with respect to the stabilizing functional $R_S(\cdot)$ in the sense that $R_S(f) = 0$ for all $f \in \mathcal{K}$. To extend the present noise-free example, the dashed line in Fig. 3 illustrates the recovered PSDF for the case when the input optical data is corrupted with 3 percent RMS Gaussian noise. For the case of noisy data, observe that a small component from $\mathcal{G}$ has been selected by the algorithm in addition to the major component from $\mathcal{K}$ (i.e., the function $h_1(r) = 1$ from $\mathcal{K}$). Note, in particular, that although a small component from $\mathcal{G}$ has been selected in the inversion, the regularization parameter selected by the generalized cross-validation procedure yields not only a relatively smooth solution but also one that provides a good approximation to the true solution.

40. Fig. 3 presents the recovered aerosol particle size distribution functions for the case of a log-normal distribution possessing a single mode at 0.5 μm. Two inversions were performed for this example: one for noise-free optical data and one for optical data that had been degraded with 5 percent RMS Gaussian noise. Again, the cross-validation procedure selected $\gamma \approx 0$ (actually $\gamma = 1.2 \times 10^{-9}$) for the noise-free inversion example. However, for the noise-corrupted example, the cross-validation procedure provided a value of 3.2×10^{-5} for γ. Observe that the inverted PSDFs compare favorably with the true size distribution with the correct mode determined at 0.5 μm. Again, note that the values of γ selected by the generalized cross-validation procedure are appropriate for both the noise-free and noise-corrupted cases since the resulting solutions obtained using these values appear to be optimally smooth—the selection of γ is neither too small resulting in an undesirable oscillatory behavior in the solution nor too large resulting in the oversmoothing of the solution with the concomitant loss of relevant detail in the PSDF (e.g., the single mode in the function). However, it is important to point out that the recovered solution for the case of noise-corrupted data exhibited (cf. Fig. 3) negative (and hence, unphysical) values at the tails of the log-normal distribution. This can probably be attributed to the the fact that the information content embodied in the optical data on the nature of the distribution in the tails is small. Consequently, features in the ends of the distribution cannot be accurately recovered using only the given noisy optical data. Furthermore, the magnitudes of the small oscillations which result in the negative values near the ends of the distribution are below the level of noise corrupting the input data.

41. Fig. 4 displays the recovered particle size distributions for the case of a bimodal log-normal distribution function with a primary mode centered at 0.5 μm and a secondary mode centered at 2.0 μm. The inversions were performed on noise-free optical data as well as data that had been corrupted at the 5 percent RMS noise level. Note that the resulting inverted solutions are good in that both modes in the distribution have been properly identified. The two peaks in the recovered PSDF fall close to the true values and the areas under these peaks approximate the actual areas quite well. The values of γ selected by the GCV procedure appear to provide stable solutions in both cases, although it should be noted that in the noise-corrupted case, the solution exhibits a small undulation at roughly the position corresponding to the relative minimum between the two modes. Fig. 5 shows the recovered PSDFs for a Junge-like distribution [25] obtained from

both noise-free optical data and data that had been degraded with 5.0 percent RMS noise. Both inversion solutions provide good estimates for the size distribution, although there are some small (minor) oscillations superimposed on the generally monotonically decreasing trend for the noise-corrupted solution. Fig. 6 displays the inverted solutions for the case of a Rosin-Rammler distribution [25]. Both noise-free optical data as well as data corrupted by Gaussian noise at the 5 percent RMS level were employed. Again, note that the agreement between the inverted and the actual size distributions is good. Indeed, the recovered PSDFs in both cases reproduce the shape of the true distribution well. This fact, coupled with the observation that the inverted results are relatively smooth, is evidence that the Lagrange multipliers γ have been properly selected by the GCV method in both the noise-free and noise-corrupted cases.

CONCLUSIONS

42. The purpose of this paper has been to develop an automatic algorithm for the "blind" reconstruction of particle size distribution functions of polydispersed aerosol clouds, composed of homogeneous spherical particles of known composition (i.e., of known refractive index), from discrete, noisy light scattering data. The algorithm is based on an alternative implementation of the popular constrained linear inversion procedure due to Philips and Twomey. This procedure tries to minimize artifacts (i.e., oscillations) in the solution by utilizing a "roughness" measure (i.e., stabilizing functional) depending on the second derivative of the PSDF. The automatic algorithm arises from the attempt to address two major problems that plague conventional implementations of the constrained linear inversion procedure, namely, the need to select an appropriate discrete approximation for the PSDF and the importance of obtaining a proper value for the Lagrange multiplier (i.e., smoothing parameter) that adequately reflects the amount of smoothing to impose on the solution in relation to the level of noise corrupting the input data.

43. To address the first of these problems, it was shown that an analytic representation for the solution of the constrained linear inversion procedure can be obtained by application of some standard techniques from the theory of Hilbert spaces. This representation provides a convenient parametrization for the PSDF and, consequently, permits the estimation of the entire size distribution rather than merely its values at nodes corresponding to some discretization of the particle size domain. This differs from the usual implementation of the constrained linear inversion procedure whereby discrete values are used both for the approximation of the PSDF and for the quadrature required to evaluate the model optical data. By constrast, in the present implementation of the procedure, the discrete values of the PSDF are utilized only for the quadrature. The second problem was addressed by application of the principle of generalized cross-validation, which provides an objective procedure for the selection of an appropriate value of the Lagrange multiplier. With an objective method for the computation of the correct value of the smoothing parameter and with the explicit representation for the PSDF, the reconstruction of the aerosol size distribution function is reduced to the problem of estimating a set of unknown coefficients whose values then determine the exact functional form of

the PSDF. This estimation problem can be efficiently resolved as the solution of a system of linear algebraic equations. It has been demonstrated using synthetic data examples that the automatic implementation of the constrained linear inversion procedure works well and requires a minimum of user intervention. These examples seem to indicate that the GCV procedure provides reliable estimates for the regularization parameter that adequately reflect the noise level corrupting the data. Indeed, the automatic algorithm when implemented with a lookup table method for the Mie kernels, provides the possibility for a rapid real-time PSDF inversion procedure using standard desktop computers.

44. In this paper, the GCV procedure was coupled with an analytic (i.e., exact) representation for the solution of the constrained linear inversion procedure in order to construct an automatic retrieval algorithm for the PSDF. However, it is important to emphasize that the GCV procedure can be coupled with other analytic representations for the PSDF in order to formulate alternative automated recovery procedures. In this connection, the eigenfunction expansion technique of Curry and Kiech [10,11] can be automated by interfacing it with the GCV procedure. It is interesting to note that the selection of the regularization parameter using the GCV procedure is closely related to the the selection using the Residual Relative Variance (RRV) procedure proposed by Curry and Kiech. Indeed, both procedures attempt to select the regularization parameter so that the computed "true" predictive root-mean-square error (i.e., the residual error between the noisy observations and the model observations computed using the true PSDF) is approximately equal to the experimental error. However, the RRV procedure requires the implementation of a somewhat awkward double iterative calculation procedure whereby a PSDF inversion computation is required for every iteration used in finding the optimal value for the regularization parameter. By contrast, the GCV procedure decouples the iterative computation for the regularization parameter from the PSDF inversion calculation.

45. The primary drawback in the application of the GCV procedure to the constrained linear inversion technique resides in the fact that it is possible to sometimes obtain negative values for the recovered PSDF. These negative values, which usually arise from high-frequency oscillations exhibited in the tails of a sharply-peaked distribution (viz., a Gibbs phenomenon) is probably the result of the small information content present in the data concerning the features in the tails of the distribution. There are two ways to overcome the negativity in these estimates: (1) by increasing the value of the regularization parameter to such a point that there are no longer negative values in the reconstructed distribution or (2) by providing a representation for the PSDF that explicitly incorporates the positivity constraints. The first remedy provides an indirect method for addressing the negativity problem and has been applied by numerous researchers [5,7,9]. Unfortunately, this method may result in an oversmoothing of the PSDF in order to satisfy the non-negativity constraints. Furthermore, the GCV procedure is no longer applicable to this case. On the other hand, the second remedy is more attractive in that the non-negativity constraints are directly applied and the GCV procedure is still applicable. In this connection, the problem needs to be solved with the constraints $\vec{a} \geq \vec{0}$ and $\vec{b} \geq \vec{0}$ imposed. Unfortunately, a closed-form solution cannot be written for this

constrained problem in contrast with the unconstrained problem and an appropriate numerical method needs to be developed.

ACKNOWLEDGEMENT

46. The author wishes to thank Dr. Y. Das and Mr. S. Mellæen for their constructive suggestions and valuable comments.

REFERENCES

1. Yamamoto, G. and Tanaka, M., "Determination of Aerosol Size Distribution From Spectral Attenuation Measurements", Appl. Opt., Vol. 8, pp. 447-453, 1969.

2. Philips, D. L., "A Technique for the Numerical Solution of Certain Integral Equations of the First Kind", J. Assoc. Comput. Mach., Vol. 9, pp. 84-97, 1962.

3. Twomey, S.,"On the Numerical Solution of Fredholm Integral Equations of the First Kind by the Inversion of the Linear System Produced by Quadrature", J. Assoc. Comput. Mach., Vol. 10, pp. 97-101, 1963.

4. Tikhonov, A. N.,"On the Solution of Incorrectly Stated Problems and a Method of Regularization", Dokl. Acad. Nauk. SSSR, Vol. 151, pp. 501-504, 1963.

5. King, M. D., Byrne, D. M., Herman, B. M. and Reagan, J. A., "Aerosol Size Distributions Obtained by Inversion of Spectral Optical Depth Measurements", J. Atmos. Sci., Vol. 35, pp. 2153-2167, 1978.

6. Dave, J. V.," Determination of Size Distribution of Spherical Polydispersions Using Scattered Radiation Data", Appl. Opt., Vol. 10, pp. 2035-2044, 1971.

7. Shaw, G. E., " Inversion of Optical Scattering and Spectral Extinction Measurements to Recover Aerosol Size Spectra", Appl. Opt., Vol. 18, pp. 988-993, 1979.

8. Herman, B. M., Browning, S. R. and Reagan, J. A., "Determination of Aerosol Size Distributions from Lidar Measurements", J. Atmos. Sci., Vol. 28, pp. 763-771, 1971.

9. King, M., "Sensitivity of Constrained Linear Inversions to the Selection of the Lagrange Multiplier", J. Atmos. Sci., Vol. 39, pp. 1356-1369, 1982.

10. Curry, B. P. and Kiech, E. L., "Two Particle Size Inversion Procedures for Interpretation of Mie Scattering Measurements (U)", AEDC-TR-83-32, November 1983, UNCLASSIFIED

11. Curry, B. P. and Kiech, E. L., "Improved Methods to Invert the Particle Size Distribution Function from Mie Scattering Measurements (U)", AEDC-TR-83-52, February 1984, UNCLASSIFIED

12. Capps, C. D., Henning, R. L. and Hess, G. M., "Analytic Inversion of Remote Sensing Data", Appl. Opt., Vol. 21, pp. 3581-3587, 1982.

13. van de Hulst, H. C., Light Scattering by Small Particles, New York: John Wiley and Sons, 1957.

14. Bohren, C. F. and Huffman, D. R., Absorption and Scattering of Light by Small Particles, New York: John Wiley and Sons, 1983.

15. Twomey, S., "The Application of Numerical Filtering for the Solution of Integral Equations Encountered in Indirect Sensing Measurements", J. Franklin Inst., Vol. 279, pp. 95–109, 1965.

16. Luenberger, D. G., Optimization by Vector Space Methods, New York: John Wiley and Sons, 1969.

17. Allen, D. M., "Mean Square Error of Prediction as a Criterion for Selecting Variables", Techonometrics, Vol. 13, pp. 469–475, 1971.

18. Wahba, G. and Wold, S., "A Completely Automatic French Curve: Fitting Spline Functions by Cross-Validation", Comm. in Statist., Vol. 4, pp. 1–17, 1975.

19. Craven, P. and Wahba, G., "Smoothing Noisy Data with Spline Functions: Estimating the Correct Degree of Smoothing by the Method of Generalized Cross-Validation", Numerische Mathematik, Vol. 31, pp. 377–403, 1979.

20. Wahba, G. "Optimal Smoothing of Density Estimates", in Classification and Clustering, pp. 423–458, ed. by J. van Ryzin, New York: Academic Press, 1976.

21. Wahba, G. and Wold, S., "Periodic Splines for Spectral Density Estimation: The Use of Cross-Validation for Determining the Correct Degree of Smoothing", Comm. in Statist., Vol. 4, pp. 125–141, 1975.

22. Dongarra, J. J., Bunch, J. R., Moler C. B. and Stewart, G. W., LINPACK Users' Guide, Philadelphia: Society for Industrial and Applied Mathematics, 1979.

23. Forsythe, G. E., Malcolm, M. A. and Moler, C. B., Computer Methods for Mathematical Computations, Englewood Cliffs, N. J.: Prentice-Hall, 1977.

24. Wiscombe, W. J., "Improved Mie scattering algorithms", Appl. Opt., Vol. 19, pp. 1505–1509, 1980.

25. Hinds, W. C., Aerosol Technology: Properties, Behavior, and Measurement of Airborne Particles, New, York: John Wiley and Sons, 1982.

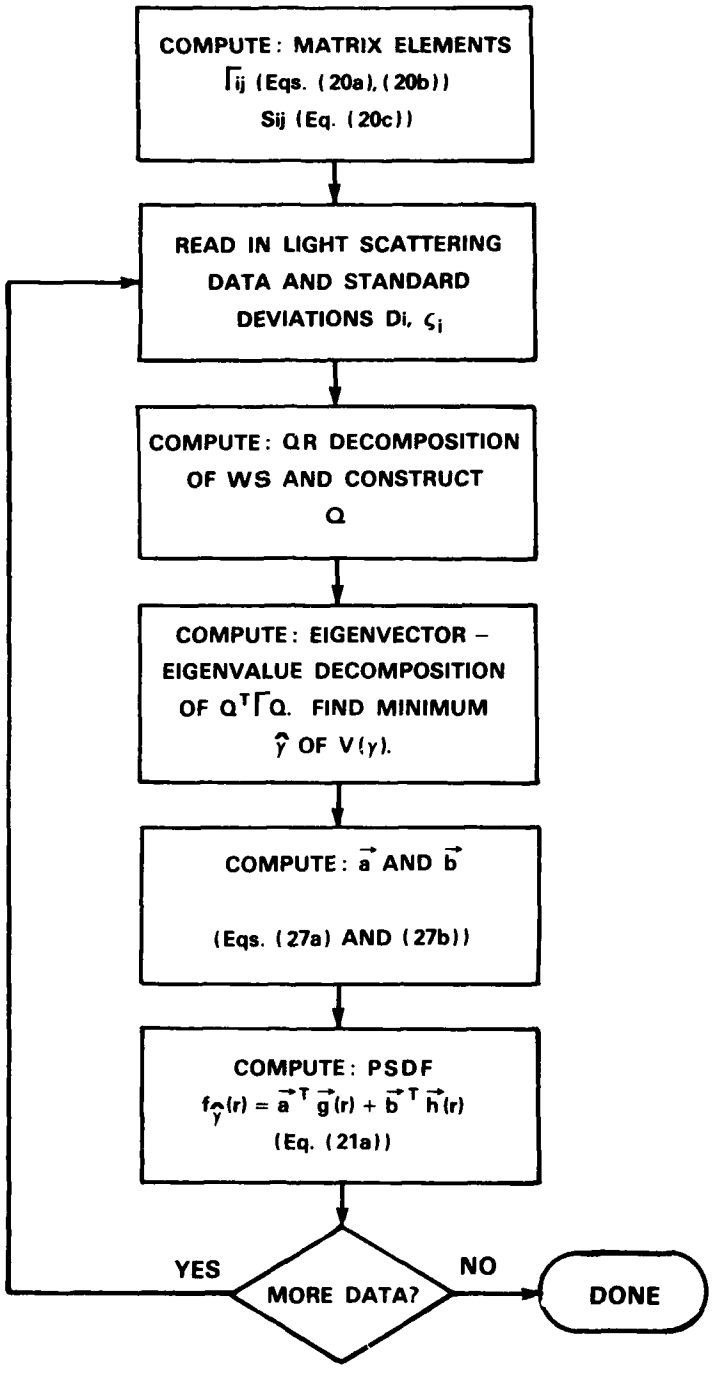

Figure 1

Flowchart of the implementation of the automatic recovery of the particle size distribution function *f(r)*.

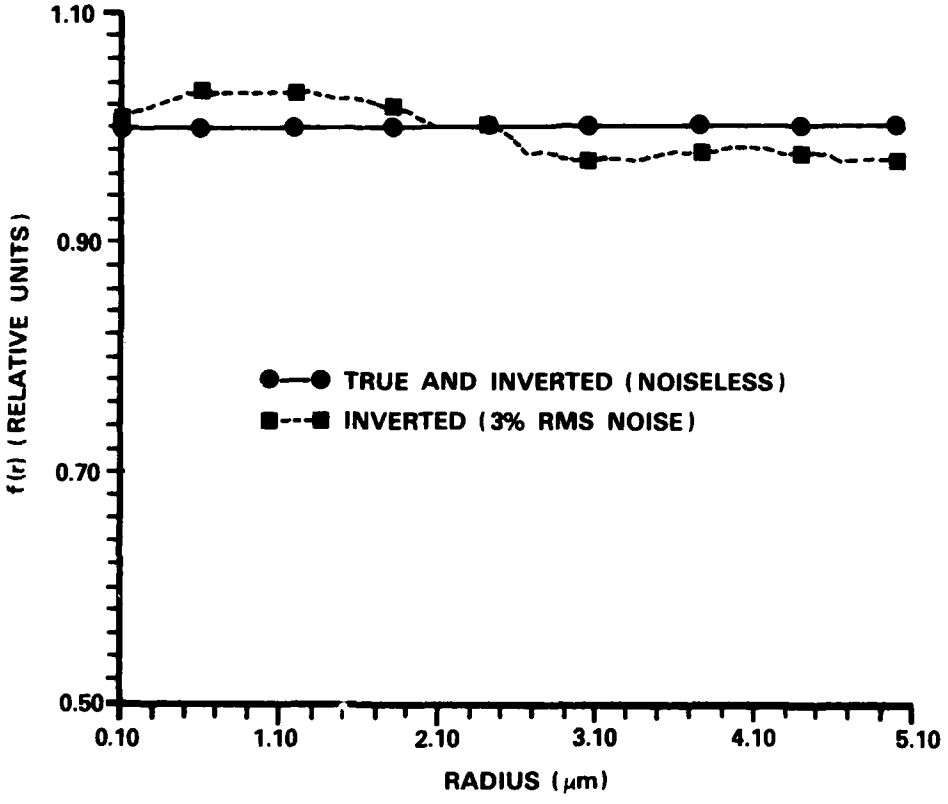

Figure 2

Recovered particle size distribution functions *f(r)* obtained for both noise-free optical data and data degraded with 3 percent RMS Gaussian noise. Also shown is the true (actual) size distribution which is uniform over the entire particle size domain Ω.

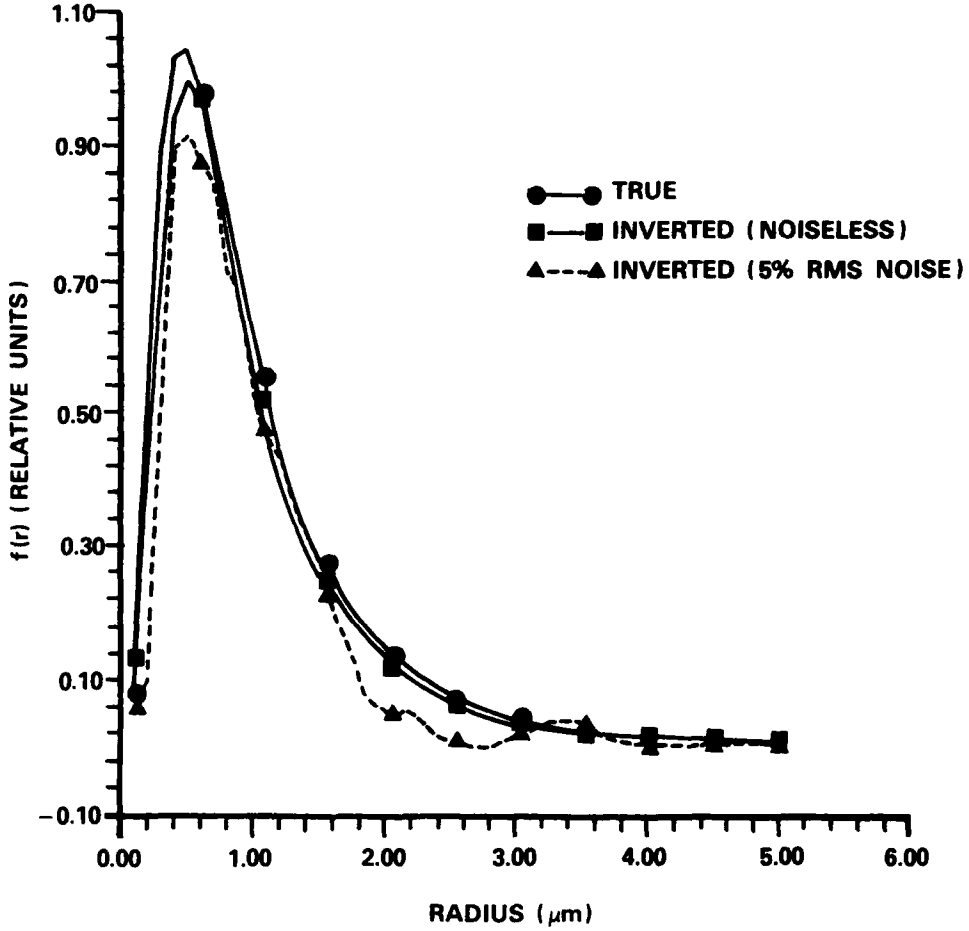

Figure 3

Recovered particle size distribution functions obtained from both noise-free optical data and data corrupted with 5 percent RMS noise. The true size distribution is a unimodal log-normal distribution with the mode centered at 0.5 μm.

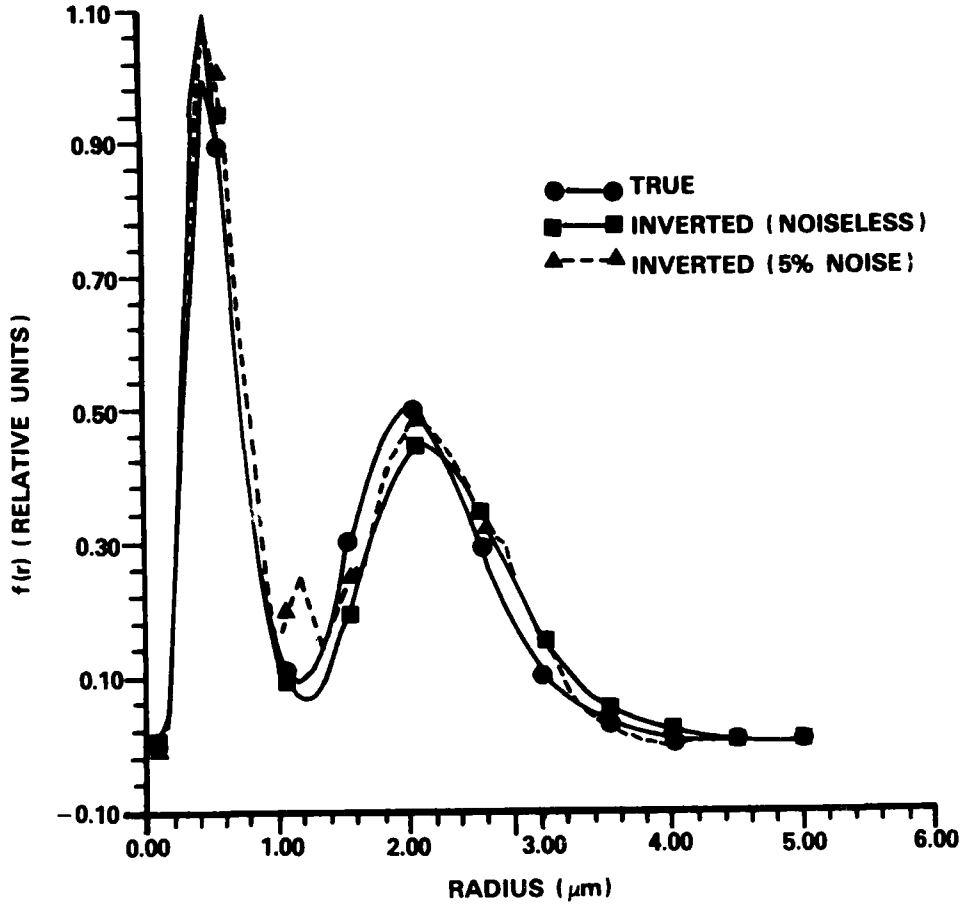

Figure 4

Recovered particle size distribution functions obtained from both noise-free optical data and data corrupted with 5 percent RMS noise. The true size distribution is a bimodal log-normal distribution with modes centered at 0.5 μm and 2.0 μm.

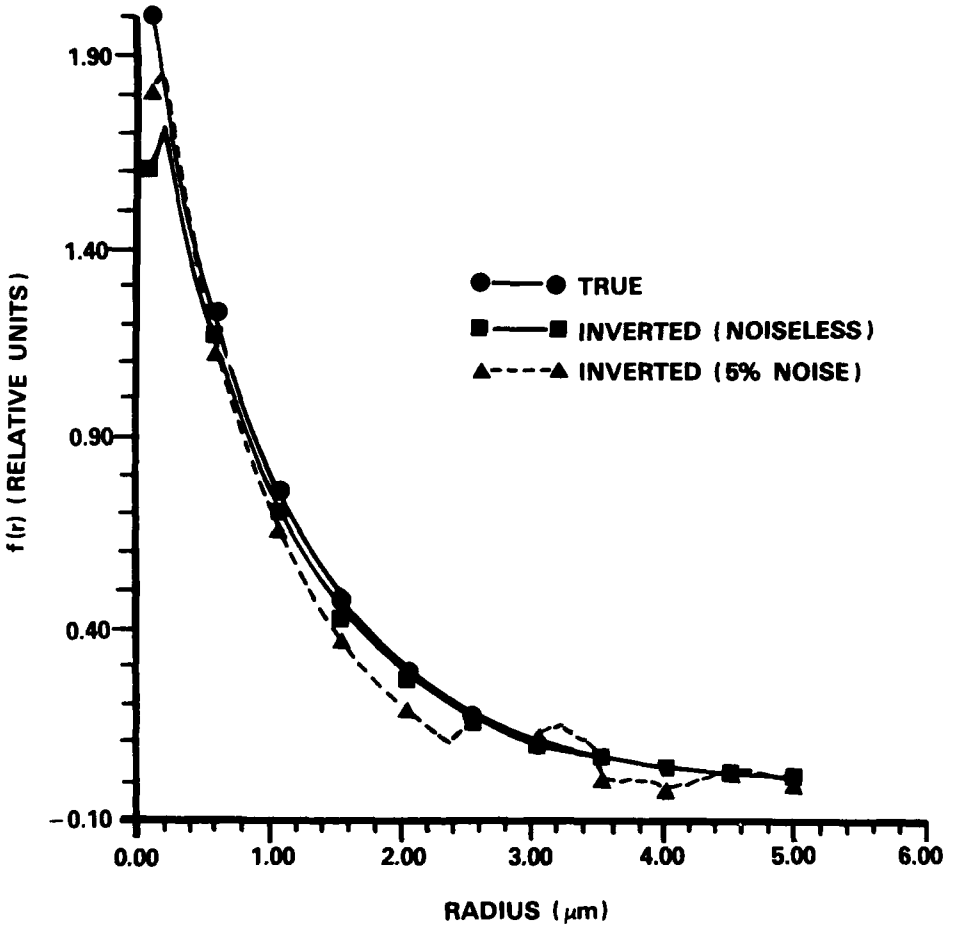

Figure 5

Recovered particle size distribution functions obtained from both noise-free optical data and data corrupted with 5 percent RMS Gaussian noise. The true size distribution is a Junge-type distribution shown by the solid line.

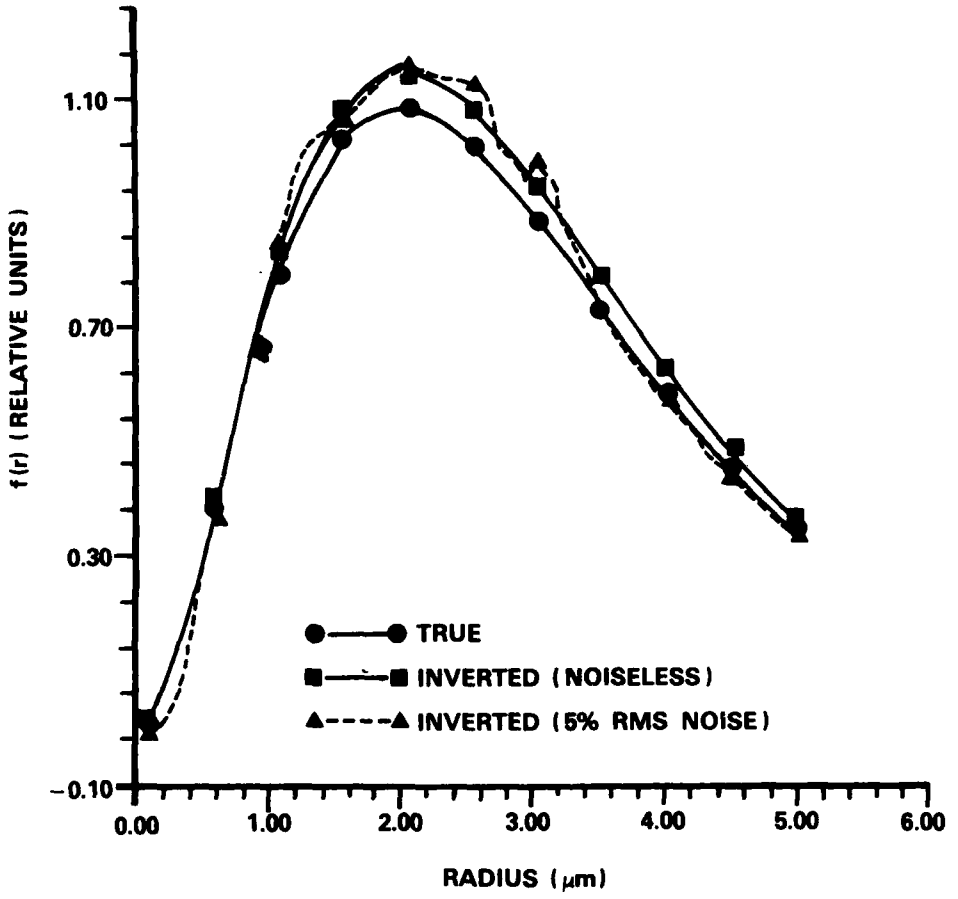

Figure 6

Recovered particle size distribution functions obtained from both noise-free optical data and data corrupted with 5 percent RMS Gaussian noise. The true size distribution is a Rosin-Rammler distribution shown by the solid line.

UNCLASSIFIED
This Sheet Security Classification

DOCUMENT CONTROL DATA - R & D
(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1 ORIGINATING ACTIVITY	2a. DOCUMENT SECURITY CLASSIFICATION
	2b. GROUP

3 DOCUMENT TITLE

INFERENCE OF PARTICLE SIZE DISTRIBUTION FUNCTION OF AEROSOL CLOUDS FROM LIGHT SCATTERING MEASUREMENTS

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (Last name, first name, middle initial)

YEE, EUGENE C.

6. DOCUMENT DATE August 1988	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS 25
8a. PROJECT OR GRANT NO.	9a. ORIGINATOR'S DOCUMENT NUMBER(S) SM 1239	
8b CONTRACT NO.	9b. OTHER DOCUMENT NO.(S) (Any other numbers that may be assigned this document)	

10. DISTRIBUTION STATEMENT

UNCLASSIFIED

11. SUPPLEMENTARY NOTES	12. SPONSORING ACTIVITY

13. ABSTRACT

Perhaps the most versatile and efficient method for inferring the particle size distribution function (PSDF) of aerosol clouds from remote light scattering measurements is the constrained linear inversion procedure due to Philips and Twomey. However, conventional numerical implementations of this procedure are subject to the following two problems: (1) an appropriate discrete approximation must be chosen for the PSDF which adequately achieves the correct balance between the conflicting requirements of resolution of detail in the PSDF and of efficiency in the computation of the solution and (2) a proper value must be selected for the Lagrange multiplier (smoothing parameter) that adequately reflects the tradeoff between the fidelity to the observed optical data and the smoothness of the solution. Consequently, an adequate recovery of the PSDF, based on the constrained linear inversion procedure, is usually achieved only after a certain amount of tedious preliminary exploratory analysis.

An alternative implementation of the constrained linear inversion procedure is presented which over-comes the problems associated with the conventional implementation. Firstly, an explicit analytical (continuous) representation for the solution of the constrained linear inversion procedure is developed which obviated the need to obtain a discrete approximation for the PSDF. Secondly, an objective procedure, based on the principle of generalized cross-validation, is utilized for the selection of the proper value for the Lagrange multiplier. Taken together, these two developments provide the basis for an objective, fully automated implementation of the constrained linear inversion technique. Numerical examples of PSDF inversions, obtained using the proposed automatic retrieval algorithm, are presented for synthetic aerosol optical extinction and scattering data.

DD FORM 1473
S/N 0102-014-6600

KEY WORDS

MIE THEORY, INVERSION, LIGHT SCATTERING

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the organization issuing the document.

2a. DOCUMENT SECURITY CLASSIFICATION: Enter the overall security classification of the document including special warning terms whenever applicable.

2b. GROUP: Enter security reclassification group number. The three groups are defined in Appendix 'M' of the DRB Security Regulations.

3. DOCUMENT TITLE: Enter the complete document title in all capital letters. Titles in all cases should be unclassified. If a sufficiently descriptive title cannot be selected without classification, show title classification with the usual one-capital-letter abbreviation in parentheses immediately following the title.

4. DESCRIPTIVE NOTES: Enter the category of document, e.g. technical report, technical note or technical letter. If appropriate, enter the type of document, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the document. Enter last name, first name, middle initial. If military, show rank. The name of the principal author is an absolute minimum requirement.

6. DOCUMENT DATE: Enter the date (month, year) of Establishment approval for publication of the document.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the document.

8a. PROJECT OR GRANT NUMBER: If appropriate, enter the applicable research and development project or grant number under which the document was written.

8b. CONTRACT NUMBER: If appropriate, enter the applicable number under which the document was written.

9a. ORIGINATOR'S DOCUMENT NUMBER(S): Enter the official document number by which the document will be identified and controlled by the originating activity. This number must be unique to this document.

9b. OTHER DOCUMENT NUMBER(S): If the document has been assigned any other document numbers (either by the originator or by the sponsor), also enter this number(s).

10. DISTRIBUTION STATEMENT: Enter any limitations on further dissemination of the document, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this document from their defence documentation center."

(2) "Announcement and dissemination of this document is not authorized without prior approval from originating activity."

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document, even though it may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall end with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (TS), (S), (C), (R), or (U).

The length of the abstract should be limited to 20 single-spaced standard typewritten lines; 7½ inches long.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a document and could be helpful in cataloging the document. Key words should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context.